CIT 5920
Recitation 7

Shenao, Sam, Tiffany, Qianyue, Shutong

Penn Engineering

Overview for Today

- Logistics
- Question 1 (4 minutes)
- Question 2 (5 minutes)
- Question 3 (3 minutes)
- Question 4 (5 minutes)
- Question 5 (4 minutes)
- Question 6 (4 minutes)

Logistics

- Midterm grades is released later today
- HW4 is extended to Wednesday, October 30th at 23:59 PM

Expectation

The expected value is the mean of the possible values a random variable can take, weighted by the probability of those outcomes.

$$E[X] = \sum_{s \in S} X(s)p(s)$$

Expectation

EXAMPLE 4.3

A men's soccer team plays soccer zero, one, or two days a week. The probability that they play zero days is .2, the probability that they play one day is .5, and the probability that they play two days is .3. Find the long-term average or expected value, μ, of the number of days per week the men's soccer team plays soccer.

To do the problem, first let the random variable X = the number of days the men's soccer team plays soccer per week. X takes on the values 0, 1, 2. Construct a PDF table adding a column $x*P(x)$, the product of the value x with the corresponding probability $P(x)$. In this column, you will multiply each x value by its probability.

x	*P(x)*	*x*P(x)*
0	.2	(0)(.2) = 0
1	.5	(1)(.5) = .5
2	.3	(2)(.3) = .6

Expectation

Add the last column $x * P(x)$ to get the expected value/mean of the random variable X.

$$E(X) = \mu = \sum xP(x) = 0 + .5 + .6 = 1.1$$

The expected value/mean is 1.1. The men's soccer team would, on the average, expect to play soccer 1.1 days per week. The number 1.1 is the long-term average or expected value if the men's soccer team plays soccer week after week after week.

Linearity of Expectation

$$E[aX + bY] = aE[X] + bE[Y]$$

where X and Y are random variables, and a and b are constants

Indicator Variables

- Take the value of 1 if the event happens and 0 if the event does not happen.

- $E(X_i) = P(X_i = 1) = \text{probability of the event happening}$

- For example, we can define an indicator variable that equals 1 if you roll an even number (2, 4, 6) and 0 if you roll an odd number.

$$\mathbb{E}(I_{\text{even}}) = P(X \in \{2, 4, 6\}) = \frac{3}{6} = \frac{1}{2}$$

Bayes' Theorem

$$P(B \mid A) = \frac{P(A \mid B)P(B)}{P(A)} = \frac{P(A \mid B)P(B)}{P(A \mid B)P(B) + P(A \mid B^c)P(B^c)}$$

Spam detection example - Given that the email contains the word "lottery", what is the probability that it is spam? Assume:

- 20% of all emails are spam: P(Spam)=0.2
- 1% of non-spam emails contain "lottery": P(Lottery | Not Spam)=0.01
- 40% of spam emails contain "lottery": P(Lottery | Spam)=0.4

$$P(\text{Spam} \mid \text{Lottery}) = \frac{P(\text{Lottery} \mid \text{Spam})P(\text{Spam})}{P(\text{Lottery} \mid \text{Spam})P(\text{Spam}) + P(\text{Lottery} \mid \text{Not Spam})P(\text{Not Spam})}$$

$$= \frac{0.4 \times 0.2}{(0.4 \times 0.2) + (0.01 \times 0.8)} \approx 0.909$$

Bernoulli Trials

Experiment with exactly two outcomes: success and failure

$$P(k) = \binom{n}{k} p^k q^{n-k}$$

where $\binom{n}{k}$ is a binomial coefficient

k is the number of successes
p is the probability of success

Question 1

The blue M&M was introduced in 1995.
Before then, the color mix in a bag of plain M&Ms was 30% Brown, 20% Yellow, 20% Red, 10% Green, 10% Orange, 10% Tan. Afterward, it was 24% Blue , 20% Green, 16% Orange, 14% Yellow, 13% Red, 13% Brown.

A friend of mine has two bags of M&Ms, and she tells me that one is from 1994 and one from 1996. She won't tell me which is which, but she gives me one M&M from each bag. One is yellow and one is green. What is the probability that the yellow M&M came from the 1994 bag?

Answer 1

Solution: Let us define the bags bag 1 and for bag 2, where bag 1 is the one that the yellow came from, and bag 2 is the one that the green came from.
We can rephrase what the question is asking in the following way: given that we picked a yellow from bag 1 and a green from bag 2, what's the probability that bag 1 is the 1994 bag?
We will define the two events:

A - bag 1 is from 1994 and bag 2 is from 1996
B - bag 1 is from 1996 and bag 2 is from 1994
YG - the event of getting a yellow from bag 1 and a green from bag 2

Then we can express the quantity that we're trying to find as $P(A|YG)$. Now by Bayes' Rule:

$$P(A|YG) = \frac{P(YG|A)P(A)}{P(YG|A)P(A) + P(YG|B)P(B)}$$

Answer 1 (cont.)

- $P(YG|A) = 0.2 \cdot 0.2$ since the chance of getting a yellow from a 1994 bag is 0.2 and the chance of getting a green from a 1996 bag is 0.2 (and they're independent).
- $P(YG|B) = 0.14 \cdot 0.1$ since the chance of getting a yellow from a 1996 bag is 0.14 and the chance of getting a green from a 1994 bag is 0.1 (and they're independent).
- $P(A) = 0.5$ and $P(B) = 0.5$ since there is a 50/50 chance each one is from 1994 vs. 1996 either way.

Therefore, the answer is

$$P(A|YG) = \frac{0.2 \cdot 0.2 \cdot 0.5}{0.2 \cdot 0.2 \cdot 0.5 + 0.14 \cdot 0.1 \cdot 0.5} = \frac{0.02}{0.02 + 0.007} = \frac{20}{27} \approx 0.741$$

Question 2

As per the rules of Blackjack, getting a Blackjack requires a total of 21 in your first two cards.
In this game, an Ace is worth 1 point or 11 points. Jacks, Queens, and Kings are all worth 10 points each, and number cards are worth their corresponding value.
You are at a casino where their rules are that you pay $10 for playing this game of Blackjack where you just get to draw 2 cards once. If you get a Blackjack, the casino will give you $30. What is your expected earning in this game?

Answer 2

Solution: First, let's find the probability of winning. There are $\binom{52}{2}$ possible ways of picking 2 cards.

To get a blackjack you must get an Ace and either a 10 or a Face card (10 or Jack, Queen, or King). You can get an Ace in 4 ways, and there are 16 ways to get a 10 or Face card (4 of the 10s, 4 of the Jacks, 4 of the Queens, and 4 of the Kings). Thus there are a total of 64 ways to get a blackjack. In the case you win, you earn a net gain of \$20. If you lose, you lose \$10. Thus the final expected gain is

$$\frac{64}{\binom{52}{2}} \cdot 20 + \left(1 - \frac{64}{\binom{52}{2}}\right) \cdot (-10) = \frac{-1890}{221} \approx -\$8.55$$

Question 3

A game is played by rolling 3 dice and betting on a number between 1 and 6.

If your number shows up once you win $1.
If your number shows up twice you win $2.
If your number shows up on all dice you win $3.
Otherwise, you win nothing.

How much should you be willing to pay to play this game?

Answer 3

Solution: There is a 1 in 6 chance of a single die having whatever number you bet on. Let us create a random variable X that takes the value 0, 1, 2, or 3 depending upon the number of dice that show the number that you bet on. We can then find the following probabilities (that you roll some number of your number):

- $P(X=0) = \left(\frac{5}{6}\right)^3$ (each one must not be your number)
- $P(X=1) = \binom{3}{1}\left(\frac{1}{6}\right)\left(\frac{5}{6}\right)^2$ (there are $\binom{3}{1}$ ways to choose the winning die, and the probability it rolled your number is $\frac{1}{6}$. The probability the other 2 were not your number is $\frac{5}{6}$)
- $P(X=2) = \binom{3}{2}\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right)$ (there are $\binom{3}{2}$ ways to choose the winning dice, and the probability they rolled your number is $\frac{1}{6}$. The probability the other one was not your number is $\frac{5}{6}$)
- $P(X=3) = \left(\frac{1}{6}\right)^3$ (each one must be your number)

Answer 3 (cont.)

Therefore, the expected earning is

$$0 \cdot P(X = 0) + 1 \cdot P(X = 1) + 2 \cdot P(X = 2) + 3 \cdot P(X = 3)$$

$$= 0 + \frac{3 \cdot 5 \cdot 5}{6^3} + 2 \cdot \frac{3 \cdot 5}{6^3} + 3 \cdot \frac{1}{6^3} = 0.5$$

Thus you should pay $0.50 to play this game in order to expect to at least break even.

Question 4

A committee of four people is drawn at random from a set of six men and three women. What is the expected gender imbalance (number of men minus number of women) in the committee? (allow the imbalance to be negative)

Answer 4

Solution: Let M be the number of men and W be the number of women, and $M - W = I$, the imbalance value. Then the values can be

- $M = 4$ and $W = 0$, which gives $I = 4$
- $M = 3$ and $W = 1$, which gives $I = 2$
- $M = 2$ and $W = 2$, which gives $I = 0$
- $M = 1$ and $W = 3$, which gives $I = -2$

Answer 4 (cont.)

Now we can find the probability of each case:

- $P(I = 4) = \dfrac{\binom{6}{4}}{\binom{9}{4}}$ (choosing 4 of the men out of 4 people total)
- $P(I = 2) = \dfrac{\binom{6}{3}\binom{3}{1}}{\binom{9}{4}}$ (choosing 3 of the men and 1 of the women out of 4 people total)
- $P(I = 0) = \dfrac{\binom{6}{2}\binom{3}{2}}{\binom{9}{4}}$ (choosing 2 of the men and 2 of the women out of 4 people total)
- $P(I = -2) = \dfrac{\binom{6}{1}\binom{3}{3}}{\binom{9}{4}}$ (choosing 1 of the men and 3 of the women out of 4 people total)

Answer 4 (cont.)

Therefore the expected value of I is

$$4 \cdot P(I = 4) + 2 \cdot P(I = 2) + 0 + (-2) \cdot P(I = -2) = \frac{10}{21} + \frac{20}{21} - \frac{2}{21} = \frac{4}{3}$$

Question 5

Exercise 5

There are n distinct balls and k distinct bins where $k < n$. The balls are picked up one by one and thrown randomly into these bins. All the balls are thrown. Each bin has the capacity to hold all n balls.

A. What is the expected number of empty bins?

B. We say we have a collision if we throw a ball into a bin that already contains another ball. Compute the expected number of collisions.

Answer 5a

Solution: Let X be the number of empty bins, and X_i be an indicator random variable that is 1 if bin i is empty, and 0 otherwise.
Then $E(X_1)$ is the expected "emptiness" of the first bin, or in other words, expected value that the first bin is empty.
Then $E(X_1) = 1 \cdot P(X1 \text{ is empty}) + 0 \cdot P(X_1 \text{ is not empty}) = P(X1 \text{ is empty}) = \left(\frac{k-1}{k}\right)^n$ since all n balls choose the other $k-1$ bins instead of X_1.
Now note that all bins have the same probability of being empty because we fairly throw balls. Therefore we have

$$E(X) = E(X_1) + E(X_2) + E(X_3) + \cdots + E(X_k) = k \cdot \left(\frac{k-1}{k}\right)^n$$

We can simplify $E(X)$ like this due to the linearity of expectation.

Answer 5b

Solution: The key to answering this question is to come up with an equation for collisions that utilizes the previous answer.

Note that an empty bin is a source of 0 collisions.

The first ball that falls into a bin is not a collision.

All subsequent balls in that bin are collisions. Therefore, we know

$$\text{number of collisions} = \text{number of balls} - \text{number of non-empty bins}$$

$$\text{number of collisions} = n - (k - \text{number of empty bins})$$

$$E(\text{number of collisions}) = n - (k - E(\text{number of empty bins}))$$

Therefore, the answer is

$$n - k + k \cdot \left(\frac{k-1}{k}\right)^n$$

Question 6

Exercise 6

Assume you have the following setup where a fair coin is tossed at each stage. There is a pot of money which begins at \$2 and gets doubled each time a head appears. The first time a tail appears, the game ends and the player wins whatever is in the pot.

What is the amount of money you would expect to win in this setup?

Answer 6

Solution: Let X be the expected winnings, and $X_1, X_2, \ldots$ be the events that we lose when game ends on turn $1, 2, \ldots$ correspondingly. If we lose, then the toss result is tail for this turn.
Also, the money we win if it ends on turn 1 is \$2, on turn 2 is $2 \cdot 2 = \$4$, turn 3 is $2 \cdot 2 \cdot 2 = \$8$, and so on. Therefore:

$$E(X) = 2 \cdot P(X_1) + 4 \cdot P(X_2) + 8 \cdot P(X_3) + \cdots + 2^i \cdot P(X_i) + \ldots$$

Since the first time we see a tail we lose, we can find:

- $P(X_1) = \frac{1}{2}$ (the first toss is a tail)
- $P(X_2) = \frac{1}{2} \cdot \frac{1}{2}$ (the first toss is a head, and the second toss is a tail)
- $P(X_3) = \left(\frac{1}{2}\right)^2 \cdot \frac{1}{2}$ (the first 2 tosses are heads, and the last toss is a tail)
- $P(X_4) = \left(\frac{1}{2}\right)^3 \cdot \frac{1}{2}$ (the first 3 tosses are heads, and the last toss is a tail)
- And so on...

Therefore

$$E(X) = 2 \cdot \frac{1}{2} + 4 \cdot \frac{1}{4} + 8 \cdot \frac{1}{8} + \cdots + 2^i \cdot \frac{1}{2^i} + \cdots = 1 + 1 + 1 + \cdots = \infty$$

See you next week!

Penn Engineering